《六祖壇經》舊版造字轉碼說明

中研院資訊所文獻處理實驗室
中研院史語所漢籍電子文獻工作小組
2008/2/29 陳建安 製作

一、 《六祖壇經》一書(六祖壇經.xml)使用舊版造字 20 個，字頻 328 次，詳如附件一。這 20 個造字中，17 個可轉成 Windows XP 能顯示的字，字頻 312 次；另外 3 個字必須轉成構字式，字頻 16 次。

二、 附件一的造字分析表說明如下：

甲、 編號：Big5 造字空間為 6217 個，編號由 1 到 6217。
乙、 造字：舊版造字
丙、 字頻(txt)：造字在「.txt」文件的出現次數
丁、 字頻(xml)：造字在「.xml」文件的出現次數
戊、 Big5：造字的 Big5 碼
己、 Unicode：造字所對應的 Unicode 碼
庚、 WinXP：造字在 Windows XP 的對應字形
辛、 構字式： Windows XP 若無對應字形，則改採用構字式
壬、 備註凡例：
1、異體字問題：新版漢籍考量到使用者檢索及使用時的便利性，將用字原則改為除專詞等特殊情形之外，一律改用標準字呈現。如編號 3800 的「杀⿰刂」係「剎」字之異體，故以「剎」字取代。又編號 1187 的「盖」係「蓋」字之簡體，亦以「蓋」字取代。

三、 《六祖壇經》中未確定字形「●」共 2 處，詳細處理內容可參考附件二的未確定字形「●」手動取代表。附件二《六祖壇經》未確定字形「●」手動取代表欄位說明如下：

甲、 檔案原文摘錄：摘錄檔案(六祖壇經.xml)中包含「●」的文句段落，其中紅字底線的部分，為需要手動取代的原文。
乙、 手動取代結果：變更過後的檔案內容，其中藍字底線的部分，為修改後的結果。

附件一、《六祖壇經》造字分析表

編號	造字	頻次(txt)	頻次(xml)	Big5	Unicode	WinXP	構字式	備註
10	辤	1	1	FA49	8FA4	辤		
1056	阤	2	2	8FD3	9641	阤		
1187	盖	4	4	90B9	76D6	盖		蓋，異體字問題。
2252	乱	11	11	9775	4E71	乱		亂，異體字問題。
3770	个	1	1	8141	4E2A	个		個，異體字問題。
3800		2	2	815F			杀⿰丆刂	剎，異體字問題。
3814	却	33	33	816D	5374	却		
4016	着	3	3	82BC	7740	着		著，異體字問題。
4028	禀	3	3	82C8	7980	禀		稟，異體字問題。
4041	竪	1	1	82D5	7AEA	竪		
4105	菓	2	2	8356	83D3	菓		
4286	从	2	2	846E	4ECE	从		從，異體字問題。
4656		13	13	86C8			爿⿱亼寸	將，異體字問題。
4738	睺	1	1	875B	777A	睺		
4768	妬	1	1	8779	59AC	妬		妒，異體字問題。
4909	弥	4	4	8869	5F25	弥		彌，異體字問題。
5095	来	61	61	89A8	6765	来		來，異體字問題。

編號	造字	頻次(txt)	頻次(xml)	Big5	Unicode	WinXP	構字式	備註
5104	礼	18	18	89B1	793C	礼		禮，異體字問題。
5434	湟	1	1	8BC1	6E7C	湟		涅，異體字問題。
5860	无	164	164	C6D3	65E0	无		

附件二、《六祖壇經》手動取代表

檔案原文摘錄	手動取代結果
吹散卷盡雲●	吹散卷盡雲霧
喜捨名為●至	喜捨名為勢至